

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2002-91470  
(P2002-91470A)

(43) 公開日 平成14年3月27日 (2002.3.27)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード(参考)
G 1 0 L	11/02	G 1 0 L 3/00	5 1 3 B 5 D 0 1 5
	15/04	3/02	3 0 1 D
	15/20	9/00	B
	21/02		D
	11/04		

審査請求 未請求 請求項の数13 O L (全 17 頁)

(21) 出願番号 特願2000-286011(P2000-286011)

(22) 出願日 平成12年9月20日 (2000.9.20)

(71) 出願人 000237592  
富士通テン株式会社  
兵庫県神戸市兵庫区御所通1丁目2番28号  
(71) 出願人 595115592  
学校法人鶴学園  
広島県広島市佐伯区三宅二丁目1-1  
(72) 発明者 大和 俊孝  
兵庫県神戸市兵庫区御所通1丁目2番28号  
富士通テン株式会社内  
(74) 代理人 100077517  
弁理士 石田 敬 (外4名)

最終頁に続く

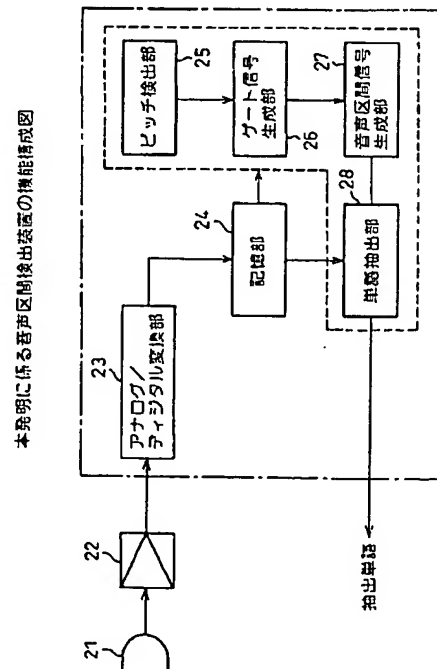
(54) 【発明の名称】 音声区間検出装置

(57) 【要約】

【課題】 促音を含む単語、あるいはサ行、ハ行音が連続する単語に対しても音声区間を確実に検出することの可能な音声区間検出装置を提供する。

【解決手段】 マイクロフォン21で検出された音声信号は、ライン増幅器22で増幅、アナログ/デジタル変換部23でデジタル化された後記憶部24に記憶される。記憶された音声信号はピッチ検出部25に取り込まれ、時間領域の処理によって音声ピッチを抽出する。ゲート信号生成部26でこの音声ピッチに基づいてゲート信号を制御し、音声区間信号生成部28でこのゲート信号に基づいて音声区間信号を制御する。音声区間信号により記憶部に記憶されている音声信号を区分することにより単語を抽出することができる。

図 2



## 【特許請求の範囲】

【請求項 1】 音声信号中に含まれる雑音を除去する前処理手段と、

前記前処理手段により雑音が除去された音声信号から音声ピッチ信号を抽出する音声ピッチ抽出手段と、

前記音声ピッチ抽出手段で抽出された音声ピッチに基づいてゲート信号を生成するゲート信号生成手段と、

前記ゲート信号生成手段に基づいて音声区間信号を生成する音声区間信号生成手段と、を具備する音声区間検出装置。

【請求項 2】 前記音声区間信号生成手段で生成された音声区間信号に基づいて前記前処理手段により雑音が除去された音声信号を複数の音声信号に区分する音声信号区分手段をさらに具備する音声区間検出装置。

【請求項 3】 前記音声ピッチ抽出手段が、前記前処理手段により雑音が除去された音声信号に対して予め定められた所定振幅以下の音声信号を除去する減算処理を施す減算処理手段と、

前記減算処理手段により減算処理された音声信号の振幅を略一定振幅に揃える揃振幅手段と、

前記揃振幅手段により略一定振幅に揃えられた音声信号から正ピーク及び前記正ピークに引き続く負ピークを検出し、前記負ピークから前記正ピークを減算して負ピークを強調した音声信号を生成する負ピーク強調手段と、前記負ピーク強調手段で負ピークの強調された音声信号を検波処理し、検波処理後の信号を微分処理する微分処理手段と、を具備する請求項 1 又は 2 に記載の音声区間検出装置。

【請求項 4】 前記減算処理手段が、前記前処理手段により雑音が除去された音声信号の正側包絡線及び負側包絡線を算出し、前記正側包絡線及び負側包絡線の差である包絡線差を算出する包絡線差算出手段と、

前記包絡線差算出手段で算出された包絡線差の予め定められた所定係数倍を乗じて減算処理閾値を算出する減算処理閾値算出手段と、

前記前処理手段により雑音が除去された音声信号の振幅が前記減算処理閾値算出手段で算出された減算処理閾値以上である場合は音声信号の振幅から減算処理閾値を減算する減算処理閾値減算手段と、を具備する請求項 3 に記載の音声区間検出装置。

【請求項 5】 前記減算処理手段が、前記前処理手段により雑音が除去された音声信号の振幅が前記減算処理閾値算出手段で算出された減算処理閾値未満である場合は音声信号の振幅を零に設定する零設定手段をさらに具備する請求項 4 に記載の音声区間検出装置。

【請求項 6】 前記揃振幅手段が、前記前処理手段により雑音が除去された音声信号の正側包絡線及び負側包絡線を算出し、前記正側包絡線及び負

側包絡線の差である包絡線差を算出する包絡線差算出手段と、

前記包絡線差算出手段で現在以前に算出された包絡線差の中の最大包絡線差を保持する最大包絡線差保持手段と、

前記最大包絡線差保持手段に保持された最大包絡線差を現在包絡線差で除して揃振幅利得を算出する揃振幅利得算出手段と、を具備する請求項 3 に記載の音声区間検出装置。

10 【請求項 7】 前記揃振幅手段が、前記揃振幅利得算出手段で算出された揃振幅利得が予め定められた所定閾値以上である場合には揃振幅利得を単位利得に設定する単位利得設定手段をさらに具備する請求項 6 に記載の音声区間検出装置。

【請求項 8】 前記ゲート信号生成手段が、前記音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値が予め定められたゲート開閾値以上となったときにゲート信号を開とするゲート信号開手段を具備する請求項 1 又は 2 に記載の音声区間検出装置。

20 【請求項 9】 前記ゲート信号生成手段が、前記ゲート信号開手段によりいったんゲート信号が開とされたときは前記音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値が前記ゲート開閾値より小である予め定められたゲート閉閾値以上であれば前記ゲート信号を開状態に維持するゲート信号開維持手段をさらに具備する請求項 8 に記載の音声区間検出装置。

【請求項 10】 前記ゲート信号生成手段が、前記音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値が、前記ゲート開閾値未満となったときに前記ゲート信号を開状態とするゲート信号閉手段をさらに具備する請求項 9 に記載の音声区間検出装置。

【請求項 11】 音声区間信号生成手段が、前記ゲート信号生成手段で生成されたゲート信号が開となった時点から予め定められた第 1 の所定期間を計時する第 1 の所定期間計時手段と、前記第 1 の所定期間計時手段による第 1 の所定期間の計時が終了した時点から予め定められた第 2 の所定期間遡及して音声区間信号を開とする音声区間信号開手段を具備する請求項 1 又は 2 に記載の音声区間検出装置。

【請求項 12】 音声区間信号生成手段が、前記ゲート信号生成手段で生成されたゲート信号が開となった時点から予め定められた第 3 の所定期間を計時する第 3 の所定期間計時手段と、前記第 3 の所定期間計時手段による第 3 の所定期間の計時が終了した時点で音声区間信号を閉とする音声区間信号閉手段をさらに具備する請求項 11 に記載の音声区間検出装置。

【請求項13】 音声区間信号生成手段が、前記第3の所定期間計時手段による第3の所定期間の計時が終了しない前に、前記音声区間信号開手段から前記第2の所定期間遡及して前記音声区間信号が開とされたときには音声区間信号を開状態に維持する音声区間信号開状態維持手段をさらに具備する請求項12に記載の音声区間検出装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は音声区間検出装置に係り、特に促音を含む単語、あるいはサ行、ハ行音が連続する単語に対しても音声区間を確実に検出することの可能な音声区間検出装置に関する。

【0002】

【従来の技術】音声認識にあつては、マイクロフォンから取り込んだ時系列信号の中から音声認識の対象となる音声区間を抽出することが必要となる。音声の短時間パワーが予め定められた閾値以上である期間を音声区間とする方法が提案されているが、不特定話者の音声から多種類の単語を認識することを目的とする場合には十分な精度を確保することは困難であった。

【0003】本出願人は既に時間領域において音声信号から声の高さであるピッチを高精度で検出することの可能なピッチ周期抽出装置及び方法をすでに提案している（特開平9-50297号公報）が、ピッチ周期に基づいて音声区間を決定することも可能である。

【0004】

【発明が解決しようとする課題】しかしながら、単語中に促音を含む単語A（例えば「壺素」）、サ行が連続する単語B（例えば「寿司屋」）、あるいはハ行が連続する単語C（例えば「皮膚科」）を対象とした場合には、単語を構成するすべての音が一つの連続した音声区間として検出されない誤検出が発生する可能性を回避できなかった。

【0005】図1は従来のピッチ周期に基づく音声区間検出結果であつて、(イ)は「単語A」を、(ロ)は「単語B」を、(ハ)は「単語C」を対象として音声区間を検出した場合を示す。いずれも上段は音声信号を、下段は音声区間を示す。この図から判明するように、「単語A」の場合には、前部の音（「壺素」にあつては「ちっ」）は音声区間内に検出されているものの、後部の音（「壺素」にあつては「そ」）は検出されていない。

【0006】「寿司屋」の場合には「すし」と「や」の間で、そして「皮膚科」の場合には「ひふ」と「か」の間で音声区間が途切れ、一つの音声区間として検出されていない。この誤検出の原因としては、以下のものが考えられる。

A：単語Aの促音「っ」に続く摩擦音「そ」、単語Bのサ行音「す」に続く摩擦音「し」はレベルが低いだけで

なく、騒音との識別が困難であるためにピッチ周期の検出自体が困難である。

【0007】B：単語に先行する気音部あるいは騒音部がなく、かつ低ピッチである場合にはピッチ周期の検出ができない。

C：単語Cの場合にはハ行音（「皮膚化」の場合は「ひふ」）と、それに続く音（「皮膚化」の場合は「か」）の間の無音期間が長い。

D：無音期間の騒音

本発明は上記課題に鑑みなされたものであつて、促音を含む単語、あるいはサ行、ハ行音が連続する単語に対しても音声区間を確実に検出することの可能な音声区間検出装置を提供することを目的とする。

【0008】

【課題を解決するための手段】第1の発明に係る音声区間検出装置にあつては、音声信号中に含まれる雑音を除去する前処理手段と、前処理手段により雑音が除去された音声信号から音声ピッチ信号を抽出する音声ピッチ抽出手段と、音声ピッチ抽出手段で抽出された音声ピッチに基づいてゲート信号を生成するゲート信号生成手段と、ゲート信号生成手段に基づいて音声区間信号を生成する音声区間信号生成手段と、を具備する。

【0009】本発明にあつては、音声信号から抽出された音声ピッチに基づいてゲート信号が制御され、このゲート信号に基づいて音声区間信号が制御される。第2の発明に係る音声区間検出装置にあつては、音声区間信号生成手段で生成された音声区間信号に基づいて前処理手段により雑音が除去された音声信号を複数の音声信号に区分する音声信号区分手段をさらに具備する。

【0010】本発明にあつては、音声信号が音声区間信号により複数の区間に区分される。第3の発明に係る音声区間検出装置にあつては、音声ピッチ抽出手段が、前処理手段により雑音が除去された音声信号に対して予め定められた所定振幅以下の音声信号を除去する減算処理を施す減算処理手段と、減算処理手段により減算処理された音声信号の振幅を略一定振幅に揃える揃振幅手段と、揃振幅手段により略一定振幅に揃えられた音声信号から正ピーク及び前記正ピークに引き続く負ピークを検出し負ピークから正ピークを減算して負ピークを強調した音声信号を生成する負ピーク強調手段と、負ピーク強調手段で負ピークの強調された音声信号を検波処理し検波処理後の信号を微分処理する微分処理手段と、を具備する。

【0011】本発明にあつては、音声ピッチが時間領域の処理により抽出される。第4の発明に係る音声区間検出装置にあつては、減算処理手段が、前処理手段により雑音が除去された音声信号の正側包絡線及び負側包絡線を算出し正側包絡線及び負側包絡線の差である包絡線差を算出する包絡線差算出手段と、包絡線差算出手段で算出された包絡線差の予め定められた所定係数倍を乗じて

減算処理閾値を算出する減算処理閾値算出手段と、前処理手段により雑音が除去された音声信号の振幅が減算処理閾値算出手段で算出された減算処理閾値以上である場合は音声信号の振幅から減算処理閾値を減算する減算処理閾値減算手段と、を具備する。

【0012】本発明にあつては、音声信号の包絡線差の所定倍が減算処理閾値とされる。第5の発明に係る音声区間検出装置にあつては、減算処理手段が、前処理手段により雑音が除去された音声信号の振幅が前記減算処理閾値算出手段で算出された減算処理閾値未満である場合は音声信号の振幅を零に設定する零設定手段をさらに具備する。

【0013】本発明にあつては、音声信号の振幅が減算処理閾値以下である場合には音声信号の振幅が零に設定される。第6の発明に係る音声区間検出装置にあつては、揃振幅手段が、前処理手段により雑音が除去された音声信号の正側包絡線及び負側包絡線を算出し正側包絡線及び負側包絡線の差である包絡線差を算出する包絡線差算出手段と、包絡線差算出手段で現在以前に算出された包絡線差の中の最大包絡線差を保持する最大包絡線差保持手段と、最大包絡線差保持手段に保持された最大包絡線差を現在包絡線差で除して揃振幅利得を算出する揃振幅利得算出手段と、を具備する。

【0014】本発明にあつては、音声信号の包絡線差に基づき揃振幅利得が決定される。第7の発明に係る音声区間検出装置にあつては、揃振幅手段が、揃振幅利得算出手段で算出された揃振幅利得が予め定められた所定閾値以上である場合には揃振幅利得を単位利得に設定する単位利得設定手段をさらに具備する。本発明にあつては、揃振幅利得が予め定められた所定閾値以上である場合は揃振幅利得は単位利得に設定される。

【0015】第8の発明に係る音声区間検出装置にあつては、ゲート信号生成手段が、音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値が予め定められたゲート開閾値以上となったときにゲート信号を開とするゲート信号開手段を具備する。本発明にあつては、予め定められた数の音声ピッチの平均値がゲート開閾値以上となったときにゲート信号が開とされる。

【0016】第9の発明に係る音声区間検出装置にあつては、ゲート信号生成手段が、ゲート信号開手段によりいったんゲート信号が開とされたときは音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値がゲート開閾値より小である予め定められたゲート閉閾値以上であればゲート信号を開状態に維持するゲート信号開維持手段をさらに具備する。

【0017】本発明にあつては、ゲート信号は連続する予め定められた数の音声ピッチの平均値がゲート閉閾値以上であればゲート信号は開状態に維持される。第10の発明に係る音声区間検出装置にあつては、ゲート信号

生成手段が、音声ピッチ抽出手段で抽出された連続する予め定められた数の音声ピッチの平均値がゲート開閾値未満となったときにゲート信号を開状態とするゲート信号開手段をさらに具備する。

【0018】本発明にあつては、ゲート信号は音声ピッチの平均値がゲート閉閾値未満となったときに閉とされる。第11の発明に係る音声区間検出装置にあつては、音声区間信号生成手段が、ゲート信号生成手段で生成されたゲート信号が開となった時点から予め定められた第1の所定期間を計時する第1の所定期間計時手段と、第1の所定期間計時手段による第1の所定期間の計時が終了した時点から予め定められた第2の所定期間遡及して音声区間信号を開とする音声区間信号開手段を具備する。

【0019】本発明にあつては、音声区間信号は、ゲート信号が第1の所定期間継続して開であるときは第1の所定期間経過時点から第2の所定期間遡及して開とされる。第12の発明に係る音声区間検出装置にあつては、音声区間信号生成手段が、第2ゲート信号生成手段で生成されたゲート信号が開となった時点から予め定められた第3の所定期間を計時する第3の所定期間計時手段と、第3の所定期間計時手段による第3の所定期間の計時が終了した時点で音声区間信号を開とする音声区間信号開手段をさらに具備する。

【0020】本発明にあつては、音声区間信号は、ゲート信号が開となった時点から第3の所定期間経過後に閉とされる。第13の発明に係る音声区間検出装置にあつては、音声区間信号生成手段が、第3の所定期間計時手段による第3の所定期間の計時が終了しない前に音声区間信号開手段から第2の所定期間遡及して音声区間信号が開とされたときには音声区間信号を開状態に維持する音声区間信号開状態維持手段をさらに具備する。

【0021】本発明にあつては、音声区間信号は、第3の所定期間と第2の所定期間が重複する場合には開状態に維持される。

【0022】

【発明の実施の形態】図2は本発明に係る音声区間検出装置の機能構成図であつて、マイクロフォン21で電気信号に変換された音声信号はライン増幅器22で増幅された後、アナログ／デジタル変換部23で音声信号は予め定められたサンプリング時間 $\Delta t$ ごとにサンプリングされてデジタル信号に変換されて、記憶部24に記憶される。

【0023】ゲート信号生成部26はピッチ検出部25で検出されたピッチに基づいてゲート信号を生成し、音声区間信号生成部27はゲート信号生成部26で生成されたゲート信号に基づいて音声区間信号を生成する。単語抽出部28は、音声区間信号生成部27で生成された音声区間信号に基づいて記憶部24に記憶されているデジタル信号を処理して音声区間に含まれる単語を抽出

して出力する。

【0024】なお、本実施例においては、アナログ／デジタル変換部23、記憶部24、ピッチ検出部25、ゲート信号生成部26、音声区間信号生成部27及び単語抽出部28は例えばパーソナルコンピュータで構成され、ピッチ検出部25、第1ゲート生成部26、第2ゲート生成部27及び単語抽出部28はソフトウェア的に構成される。

【0025】図3はアナログ／デジタル変換部23及び記憶部24で実行される音声サンプリングルーチンのフローチャートであって、サンプリング時間 $\Delta t$ ごとに割り込み処理として実行される。まず、ステップ30でアナログ／デジタル変換部23でサンプリングされた音声信号Vを取り込み、ステップ31で音声信号Vに対して前処理を行うが前処理の内容の詳細は後述する。

【0026】ステップ32で記憶部23の記憶順序を示すインデックスiを"1"に設定し、ステップ33から35で既に記憶部23に記憶されている音声信号X(i)を以下の処理により順送りする。

$X(i+1) \leftarrow X(i)$

順送りが完了すると、今読み込んだ音声信号Vを記憶部23の最先番地X(1)に記憶してこのルーチンを終了する。

【0027】図4はステップ31で実行される前処理ルーチンの詳細フローチャートであって、ステップ310でデジタル信号に対して高周波雑音除去処理を実行する。この処理には、例えば遮断周波数が4KHzであり遮断特性が18dB/octであるローパスフィルタが使用される。ステップ311では高周波雑音が除去されたデジタル信号に対して低周波雑音除去処理を実行する。この処理には、例えば遮断周波数が300Hzであり遮断特性が18dB/octであるハイパスフィルタが使用される。

【0028】なお、上記実施例においては高周波雑音除去処理及び低周波雑音除去処理をソフトウェアで行っているが、ライン増幅器22内にハードウェア的なフィルタを組み込んでよい。図5はピッチ検出部25で実行されるピッチ検出ルーチンの詳細フローチャートであって、ステップ50で記憶部23に記憶されている音声信号X(i)を読み込む。

【0029】そして、ステップ51で減算処理を、ステップ52でAGC処理を、ステップ53でピーク検出処理を実行する。さらに、ステップ54で極値検出クランプ処理を、ステップ55でピッチ周期検出処理を実行してこのルーチンを終了する。なお、ステップ51～55の処理については、以下に詳述する。

【0030】図6はピッチ検出ルーチンのステップ51で実行される減算処理ルーチンのフローチャートであって、音声信号の振幅を一定に揃えるAGC処理において微小な雑音成分までもAGC処理されて増幅されること

を防止するために所定振幅以下の成分を除去することを目的とする。まずステップ51aで包絡値差 $\Delta E$ を算出するが詳細は図7で説明する。

【0031】ステップ51bで包絡値差 $\Delta E$ が予め定められた振幅除去の閾値r未満であるかを判定し、肯定判定されたとき、即ち包絡値差 $\Delta E$ が閾値r未満であるときはステップ51cで音声信号X(i)を"0"に設定してステップ51dに進む。なお、ステップ51bで否定判定されたとき、即ち包絡値差 $\Delta E$ がしきい値r以上であるときは直接ステップ51dに進む。

【0032】ステップ51dでは今回の正側包絡値 $E_p$ が前回の正側包絡値 $E_{pb}$ より大であるかを判定する。ステップ51dで肯定判定されたとき、即ち今回の正側包絡値 $E_p$ が前回の正側包絡値 $E_{pb}$ より大であり正側包絡値が増加しているときは、ステップ51eでインデックスsを"1"に設定してステップ51gに進む。

【0033】逆にステップ51dで否定判定されたとき、即ち今回の正側包絡値 $E_p$ が前回の正側包絡値 $E_{pb}$ より小であり正側包絡値が減少しているときは、ステップ51fでインデックスsを"0"に設定してステップ51gに進む。ステップ51gではインデックスsの前回値 $s_b$ が"1"かつ今回のインデックスsが"0"である、即ち正側のピークが検出されたかを検出する。

【0034】ステップ51gで肯定判定されたとき、即ち正側ピークが検出されたときは、ステップ51hで減算処理の閾値bcを次式を用いて算出した後ステップ51iに進む。

$bc \leftarrow a * \Delta E$

ここでaは予め定められた所定値であり、本発明に係る音声区間検出装置を自動車車室内で使用する場合には一定値"0.05"とすることができる。

【0035】逆にステップ51gで否定判定されたとき、即ち正側ピークが検出されなかったときは、直接ステップ51iに進む。ステップ51iでは音声信号X(i)が減算処理の閾値bc以上であるか、即ち音声信号X(i)の振幅が大であるかを判定する。ステップ51iで肯定判定されたとき、即ち音声信号X(i)の振幅が閾値bc以上であるときは、ステップ51jで音声信号X(i)から減算処理の閾値bcを減算した値を減算処理後音声信号 $X_s(i)$ に設定してステップ51iに進む。

【0036】 $X_s(i) \leftarrow X(i) - bc$   
一方、ステップ51iで否定判定されたとき、即ち音声信号X(i)の振幅が閾値bc未満であるときはステップ51kで $X_s(i)$ を零に設定してステップ51iに進む。なお、ステップ51kの処理を省略してステップ51iで否定判定されたときは直接ステップ51iに進むようにしてもよい。

【0037】最後に、ステップ51iで前回の正側包絡値 $E_{pb}$ 、前回の負側包絡値 $E_{ab}$ 及び前回のインデックス

$s_b$  を更新してこのルーチンを終了する。

$E_{pb} \leftarrow E_p$

$E_{mb} \leftarrow E_m$

$s_b \leftarrow s$

図7は減算処理ルーチンのステップ51aで実行される包絡値差算出ルーチンのフローチャートであって、ステップa1で今回正側包絡値 $E_p$ を次式により算出する。

【0038】

$$E_p = E_{pb} \cdot \exp \{-1 / (\tau \cdot f_s)\}$$

ここで $\tau$ は時定数

$f_s$ はサンプリング周波数

ステップa2で今回負側包絡値 $E_m$ を次式により算出する。

$$E_m = E_{mb} \cdot \exp \{-1 / (\tau \cdot f_s)\}$$

ステップa3で減算処理後の音声信号 $X_s(i)$ とステップa1で演算された今回正側包絡値 $E_p$ の最大値を改めて今回正側包絡値 $E_p$ に置き換える。

【0039】ステップa4で減算処理後の音声信号 $X_s(i)$ とステップ71で算出された今回負側包絡値 $E_m$ の最小値を改めて今回負側包絡値 $E_m$ に置き換える。ステップa5で次式により包絡値差 $\Delta E$ を算出して、このルーチンを終了する。

$$\Delta E = E_p - E_m$$

図8は減算処理の効果に説明図であって、(イ)は減算処理前の音声信号を、(ロ)は減算処理後の音声信号を示す。この図から、減算処理により微小な雑音が除去されていることが理解される。

【0040】図9はピッチ検出ルーチンのステップ52で実行されるAGC処理ルーチンのフローチャートであって、減算処理後の音声信号 $X_s(i)$ の振幅を一定に揃えることを目的とする。まず、ステップ52aで最大包絡値差 $\Delta E_{max}$ の初期値を"0"に設定し、ステップ52bで図7に示す包絡値差算出ルーチンを実行して包絡値差 $\Delta E$ を算出する。ただし、この場合は包絡値差算出ルーチンのステップa3及びa4の $X(i)$ を $X_s(i)$ とすることはいうまでもない。

【0041】次にステップ52cで、

$$X_s(i-2) < X_s(i-1)$$

かつ

$$X_s(i) < X_s(i-1)$$

かつ

$$X(i-1)s > 0$$

であるか、即ち $\Delta t$ 前にサンプリングされた減算処理後の音声信号 $X_s(i-1)$ が正のピークであるかを判定する。

【0042】ステップ52cで肯定判定されたとき、即ち減算処理後の音声信号 $X_s(i-1)$ が正のピークであるときは、ステップ52dで包絡値差 $\Delta E$ とそれ以前に決定された最大包絡値差 $\Delta E_{max}$ の最大値を最大包絡値差 $\Delta E_{max}$ に設定し直して、最大包絡値差 $\Delta E_{max}$ を

更新してステップ52eに進む。なお、ステップ52cで否定判定されたとき、即ち音声信号 $X_s(i-1)$ が正のピークでないときは直接ステップ52eに進む。

【0043】ステップ52eではステップ52bで算出した包絡値差 $\Delta E$ が"0"であるかを判定する。そして、否定判定されたとき、即ち $\Delta E$ が"0"でないときはステップ52fでゲイン $G$ を $\Delta E_{max} / \Delta E$ に設定する。次にステップ52gでゲイン $G$ が予め定められた閾値 $\beta$ (例えば10)以上であるかを判定し、肯定判定されたときはステップ52hでゲイン $G$ を"1"に設定してステップ52iに進む。なお、ステップ52gの判断を省略して、ステップ52fからステップ52iに直接進むようにしてもよい。

【0044】逆にステップ52gで否定判定されたとき、即ちゲイン $G$ が予め定められた閾値 $\beta$ 未満であるときは直接ステップ52iに進む。なお、ステップ52eで肯定判定されたとき、即ち $\Delta E$ が"0"である時もステップ52hでゲイン $G$ を"1"に設定してステップ52iに進む。最後にステップ52iで減算処理後の音声信号 $X_s(i-1)$ にゲイン $G$ を乗算してAGC処理後の音声信号 $X_g(i-1)$ を算出してこのルーチンを終了する。

【0045】

$$X_g(i-1) \leftarrow G * X_s(i-1)$$

図10はAGC処理の効果の説明図であって、(イ)はAGC処理前の音声信号を、(ロ)はAGC処理後の音声信号を示す。即ち、(イ)のように音声波形の振幅が急激に変化する場合には後述のピッチ周期検出において誤検出の発生を回避できない。そこで、AGC処理により音声波形をほぼ一定振幅に揃えることにより、誤検出の発生を防止することが可能となる。

【0046】図11はピッチ検出ルーチンのステップ53で実行されるピーク検出処理ルーチンの詳細フローチャートであって、ステップ53aでAGC処理後の音声信号に正ピークが検出されたかを判定する。即ち、以下の条件が満足されたときに $X_g(i-2)$ が正ピークであると判定する。

$$X_g(i-3) < X_g(i-2) \quad \text{かつ} \quad X_g(i-1) < X_g(i-2) \quad \text{かつ} \quad 0 < X_g(i-2)$$

ステップ53aで肯定判定されたとき、即ちAGC処理後の音声信号に正ピークが検出されたときはステップ53bでピーク値 $X_g(i-2)$ をPとして記憶してこのルーチンを終了する。

【0047】ステップ53aで否定判定されたとき、即ちAGC処理後の音声信号に正ピークが検出されないときは直接このルーチンを終了する。図12はピッチ検出ルーチンのステップ54で実行される極値検出・クランプ処理ルーチンの詳細フローチャートであって、ステップ324aでAGC処理後の音声信号 $X_g$ に負のピークが検出されたかを判定する。即ち、以下の条件が満足さ

れたときに  $X_G(i-2)$  が負ピークであると判定する。

【0048】  $X_G(i-3) > X_G(i-2)$  かつ  $X_G(i-1) > X_G(i-2)$  かつ  $0 > X_G(i-2)$

ステップ54aで肯定判定されたとき、即ちAGC処理後の音声信号に負のピークが検出されたときは、ステップ54bでAGC処理後の音声信号  $X_G(i-2)$  からピーク値  $P$  を減算して、負ピークを強調したクランプ処理後の音声信号  $X_C(i-2)$  を算出してこのルーチンを終了する。

【0049】

$$X_C(i-2) \leftarrow X_G(i-2) - P$$

ステップ54aで否定判定されたとき、即ちAGC処理後の音声信号に負のピークが検出されないときは、ステップ54cでAGC処理後の音声信号  $X_G(i-2)$  をクランプ処理後の音声信号  $X_C(i-2)$  としてこのルーチンを終了する。

【0050】  $X_C(i-2) \leftarrow X_G(i-2)$

図13はピッチ検出ルーチンのステップ55で実行されるピッチ周期検出処理ルーチンの詳細フローチャートであって、ステップ55aにおいて検波後出力  $X_D(i-3)$  を次式により算出する。

$$X_D(i-3) \leftarrow E \cdot \exp(-\Delta t / \tau)$$

ここで  $\Delta t$  はサンプリング時間、 $\tau$  は予め定められた時定数である。

【0051】なお、 $E$ については後述する。ステップ55bでクランプ処理後の音声信号  $X_C(i-3)$  の絶対値が検波後出力  $X_D(i-3)$  の絶対値より大であるかを判定する。ステップ55bで否定判定されたとき、即ち  $X_C(i-3)$  の絶対値が  $X_D(i-3)$  の絶対値以下であるときは、ステップ55cで検波後出力  $X_D(i-3)$  を  $E$  に設定してステップ55fに進む。

【0052】ステップ55bで肯定判定されたとき、即ち  $X_C(i-3)$  の絶対値が  $X_D(i-3)$  の絶対値より大であるときは、ステップ55dでクランプ処理後の音声信号に負のピークが存在するかを判定する。即ち、以下の条件が満足されたときに  $X_C(i-3)$  が負ピークであると判定する。

$$X_C(i-4) > X_C(i-3) \quad \text{かつ} \quad X_C(i-2) > X_C(i-3) \quad \text{かつ} \quad 0 > X_C(i-3)$$

ステップ55dで肯定判定されたとき、即ちクランプ処理後の音声信号に負のピークが検出されたときはステップ55eで負ピーク値  $X_C(i-3)$  を  $E$  に設定してステップ55fに進む。なお、ステップ55dで否定判定されたとき、即ちクランプ処理後の音声信号に負のピークが検出されないときはステップ55cに進む。

【0053】ステップ55fでは  $E$  として記憶されていた値を検波後信号  $X_D(i-3)$  に設定し、ステップ55gで次式により検波後信号変化  $\Delta X_D$  を算出する。

$$\Delta X_D \leftarrow X_D(i-3) - X_D(i-4)$$

ステップ55hで検波後信号変化  $\Delta X_D$  の絶対値が予め定められた閾値  $\gamma$  以上であるかを判定する。

【0054】ステップ55hで肯定判定されたとき、即ち検波後出力が急減したときは、ステップ55iで音声ピッチ信号  $X_P(i-3)$  を "1" に設定してこのルーチンを終了する。逆にステップ55hで否定判定されたとき、即ち検波後出力が急減していないときは、ステップ55jで音声ピッチ信号  $X_P(i-3)$  を "0" に設定してこのルーチンを終了する。

【0055】図14及び15は本発明で適用されるピッチ周期検出方法の説明図(1/2)及び(2/2)である。図14の(イ)はクランプ処理後の音声信号を、また(ロ)及び(ハ)は該当部分の拡大音声信号を示し、横軸は時間を、縦軸は振幅を表す。即ち、クランプ処理後の音声信号が負ピークを起点とする包絡線の内側にあるとき(ロ)は包絡線を維持し、外側にあるとき(ハ)はクランプ処理後の音声信号を検波後出力とする。

【0056】図15の(ニ)は検波後信号を、(ホ)は音声ピッチ信号の波形図であって、時刻  $t_2$ 、 $t_4$  及び  $t_6$  でピッチパルスが検出されていることを示す。図16は第1ゲート生成部26で実行される第1のゲート信号生成ルーチンのフローチャートであって、ステップ160で音声ピッチ信号  $X_P(i-3)$  が "1" であり、かつ直前に音声ピッチ信号が "1" であった時刻を示すインデックス  $j$  が  $(i-3)$  と等しくないかを判断する。

【0057】ステップ160で否定判定されたとき、即ち音声ピッチ信号  $X_P(i-3)$  が "1" でないか、又は  $j$  が  $(i-3)$  と等しいときは直接このルーチンを終了する。ステップ160で肯定判定されたとき、即ち音声ピッチ信号  $X_P(i-3)$  が "1" であり、かつインデックス  $j$  が  $(i-3)$  と等しくないときはステップ161に進み、次式によってピッチ周波数  $f$  を算出する。

【0058】

$$f(i-3) = f_s / |i(i-3) - j|$$

ここで、 $f_s$  はサンプリング周波数で  $1/\Delta t$  に等しい。ステップ162でピッチ周波数  $f$  が予め定められた最高周波数 500 Hz 以上であるかを判定し、最高周波数以上であればステップ163でピッチ周波数  $f$  を "0" に設定してステップ164に進む。なお、ステップ162で否定判定されたときは直接ステップ164に進む。

【0059】ステップ164で直前に音声ピッチ信号が "1" であった時刻を示すインデックス  $j$  を  $(i-3)$  で更新する。次に、ステップ165で次式によりピッチ周波数を更新した後、平均ピッチ周波数  $f_m$  を算出する。なお、本実施形態においては3つのピッチ周波数の算術平均により平均ピッチ周波数を算出しているが、



使用するピッチ周波数の数は3つに限定されない。又平均ピッチ周波数の算出方法も算術平均に限定されず、重み付け平均、移動平均等の方法により算出してもよい。

【0060】

$f_3 \leftarrow f_2$

$f_2 \leftarrow f_1$

$f_1 \leftarrow f(i-3)$

$f_m = (f_3 + f_2 + f_1) / 3$

そして、ステップ166で平均ピッチ周波数  $f_m$  が予め定められた第1の閾値  $T_{h1}$  (例えば200Hz) 以上であるかを判定する。

【0061】ステップ166で肯定判定されたとき、即ち平均ピッチ周波数  $f_m$  が第1の閾値  $T_{h1}$  以上であるときは、音声区間が始まったものとしてステップ167で第1ゲート信号  $g_1$  を "1" に設定してこのルーチンを終了する。逆に、ステップ166で否定判定されたとき、即ち平均ピッチ周波数  $f_m$  が第1の閾値  $T_{h1}$  未満であるときは、ステップ168で平均ピッチ周波数  $f_m$  が予め定められた第2の閾値  $T_{h2}$  (例えば80Hz) 以上であるかを判定する。

【0062】ステップ168で肯定判定されたとき、即ち平均ピッチ周波数  $f_m$  が第2の閾値  $T_{h2}$  以上であるときは、音声区間が継続しているものとしてステップ167に進みゲート信号  $g_1$  を "1" に維持してこのルーチンを終了する。逆に、ステップ168で否定判定されたとき、即ち平均ピッチ周波数  $f_m$  が第2の閾値  $T_{h2}$  未満であるときは、音声区間が終了したものとステップ169に進みゲート信号  $g_1$  を "0" にリセットしてこのルーチンを終了する。

【0063】図17はゲート信号の生成方法の説明図であって、(イ)はピッチ周波数を、(ロ)はゲート信号  $g_1$  を示す。そして、(イ)の黒丸は各時刻におけるピッチ周波数  $f$  を表す。即ち、連続する3つの平均ピッチ周波数が第1の閾値  $T_{h1}$  (200Hz) 以上となった時点でゲート信号  $g_1$  が "1"、即ち開となる。

【0064】そして連続する3つの平均ピッチ周波数が第2の閾値  $T_{h2}$  (80Hz) 以上を維持している間はゲート信号  $g_1$  は開を維持し、連続する3つの平均ピッチ周波数が第2の閾値  $T_{h2}$  (80Hz) 未満となったときにゲート信号  $g_1$  は "0"、即ち閉となる。図18は音声信号処理例であって、(イ)は対象の音声信号  $V$  を前処理ルーチンで遮断周波数が300Hzである高周波通過フィルタでろ波して低周波雑音を除去した音声信号  $X$  を示す。

【0065】(ロ)はAGC処理ルーチンによるAGC処理後の音声信号  $X_0$  の波形であって、所定の振幅以上の成分の振幅がほぼ一定に形成されている。(ハ)はピッチ周期検出処理ルーチンによる検波処理後の信号  $X_0$  を示し、(ニ)は第1のゲート信号生成ルーチンのステ

ップ341で算出されたピッチ周波数  $f$  を示す。

【0066】さらに、(ホ)は第1のゲート信号生成ルーチンで生成されたゲート信号  $g_1$  を示す。この図から理解できるように音声信号の存在期間とゲート信号  $g_1$  が開である期間は一致するものの、音声途絶えた後に雑音が発生した場合には雑音に起因したピッチ周波数(ニ)の○印が発生しゲート信号  $g_1$  の閉タイミングが遅れてしまう。

【0067】図19は第2のゲート信号生成ルーチンのフローチャートであって、第1のゲート信号生成ルーチンに対してステップ190~193が追加され、上記課題を解決することを目的とする。即ち、ステップ190で次式により直前に音声ピッチ信号  $X_p(i-3)$  が "-1" であった時刻を示すインデックス  $j$  から  $(i-3)$  までの経過時間  $D_t$  を算出する。

【0068】 $D_t \leftarrow |(i-3) - j| / f_s$

次にステップ191で経過時間  $D_t$  が予め定められた閾値時間  $D_{tth}$  (例えば0.025秒) 以上、かつゲート信号  $g_1$  が "1" (即ちゲートが開) であるかを判定する。そしてステップ191で肯定判定されたとき、即ちゲートが開であり、かつ最後に "-1" である音声ピッチ信号が検出されてから25ミリ秒以上が経過したときはステップ193で修正ゲート信号  $g_1$  を "0" にしてゲートを閉とするとともに、インデックス  $j$  を更新し、 $f_2$  及び  $f_3$  をリセットしてこのルーチンを終了する。

【0069】逆にステップ192で否定判定されたとき、即ちゲートが開であるか、あるいは最後に "-1" である音声ピッチ信号が検出されてから25ミリ秒が経過していないときはステップ194で図16に示す第1のゲート信号生成ルーチンを実行してこのルーチンを終了する。なお、上記実施形態において閾値時間  $\Delta t_{th}$  を25ミリ秒としたのは、25ミリ秒以上は周波数40Hz以下に対応するが、人間の声のピッチ周波数が40Hz以下となることは考え難いからである。

【0070】第2のゲート信号生成ルーチンで生成された修正ゲート信号を図18の(ヘ)に示すが、雑音に起因したピッチ周波数(ニ)の○印の影響を受けずに修正ゲートが開となることが判る。上記の修正ゲートを使用することにより正確に音声区間を検出することが可能であるが、さらに以下の課題を解決することにより一層正確な音声区間の検出が可能となる。

【0071】1. 3つのピッチ周波数の平均値が第1の閾値  $T_{h1}$  以上となったときにゲートを開としているので、開タイミングが遅れがちとなる。

2. 単発的な大振幅の雑音と音声信号を区別できない。

3. 気音と雑音を区別できない。

4. 促音は振幅が小であるため、促音を検出できない。

【0072】そこで本発明においては、ゲート信号(修正ゲート信号も含む)によって以下のように制御される音声区間信号を導入することにより上記課題を解決す



る。即ち上記 1、2 及び 3 を解決するために、ゲート信号が第 1 の所定期間（例えば 50 ミリ秒）以上開状態を維持している場合に、現在時点から第 2 の所定期間（例えば 100 ミリ秒）遡及して音声区間信号を開とする。

【0073】上記 4 を解決するために、ゲート信号が開となった時点から第 3 の所定期間（例えば 150 ミリ秒）は音声区間信号を開状態に維持する。図 20 は音声区間信号生成部 27 で実行される音声区間信号生成ルーチンのフローチャートであって、ステップ 200 で前回演算されたゲート信号  $g_{1b}$  が "0" であるか、即ちゲートが閉であったかが判定される。

【0074】ステップ 200 で肯定判定されたとき、即ちゲートが閉であったときはステップ 201 で今回演算されたゲート信号  $g_1$  が "0" であるか、即ちゲートが閉を維持しているかを判定する。ステップ 201 で肯定判定されたとき、即ちゲートが閉を維持しているときはステップ 202 で閉維持処理を実行してステップ 207 に進む。

【0075】ステップ 201 で否定判定されたとき、即ちゲートが開に移行したときはステップ 203 で開処理を実行してステップ 207 に進む。一方ステップ 200 で否定判定されたとき、即ちゲートが開であったときはステップ 204 で今回演算されたゲート信号  $g_1$  が "1" であるか、即ちゲートが開を維持しているかを判定する。

【0076】ステップ 204 で肯定判定されたとき、即ちゲートが開を維持しているときはステップ 205 で閉維持処理を実行してステップ 207 に進む。ステップ 204 で否定判定されたとき、即ちゲートが閉に移行したときはステップ 206 で閉処理を実行してステップ 207 に進む。ステップ 207 で音声区間信号を出力し、ステップ 208 では前回演算されたゲート信号  $g_{1b}$  を今回演算されたゲート信号  $g_1$  で更新してこのルーチンを終了する。

【0077】図 21 は音声区間信号生成ルーチンのステップ 202 で実行される閉維持処理ルーチンのフローチャートであって、ステップ 2a でゲート信号  $g_1$  が閉状態を継続している時間である閉継続時間  $t_{c0}$  にサンプリング時間  $\Delta t$  を加算する。ステップ 2b で閉継続時間  $t_{c0}$  が第 3 の所定期間である 150 ミリ秒以上となったかを判定する。

【0078】ステップ 2b で肯定判定されたとき、即ちゲート信号  $g_1$  が閉となってから 150 ミリ秒を経過したときは、ステップ 2c で処理時刻  $(i-3)$  における音声区間信号  $g_2$  を "1" に設定してこのルーチンを終了する。逆にステップ 2b で否定判定されたとき、即ちゲート信号  $g_1$  が閉となってから 150 ミリ秒を経過していないときは、ステップ 2d で処理時刻を示すインデックスが  $(i-3)$  であるときの第 2 のゲート信号である  $g_2(i-3)$  を "0" に設定してこのルーチを終

了する。

【0079】図 22 は音声区間信号生成ルーチンのステップ 203 で実行される開処理ルーチンのフローチャートであって、ステップ 3a で前回演算されたゲート信号  $g_{1b}$  を "1" に設定する。ステップ 3b で閉継続時間  $t_{c0}$  を "0" に復帰し、ステップ 3c で処理時刻を示すインデックスが  $(i-3)$  であるときの音声区間信号である  $g_2(i-3)$  を "1" に設定してこのルーチンを終了する。

10 【0080】図 23 は音声区間信号生成ルーチンのステップ 205 で実行される閉維持処理ルーチンのフローチャートであって、ステップ 5a でゲート信号  $g_1$  が開状態を継続している時間である閉継続時間  $t_{c0}$  にサンプリング時間  $\Delta t$  を加算する。ステップ 5b で閉継続時間  $t_{c0}$  が第 1 の所定期間である 50 ミリ秒以上となったかを判定する。

【0081】ステップ 5b で否定判定されたとき、即ちゲート信号  $g_1$  が開となってから 50 ミリ秒を経過していないときはステップ 5c で処理時刻を示すインデックスが  $(i-3)$  であるときの音声区間信号である  $g_2(i-3)$  を "0" に設定してこのルーチンを終了する。ステップ 5b で否定判定されたとき、即ちゲート信号  $g_1$  が開となってから 50 ミリ秒を経過したときは、ステップ 5d で処理時刻から第 2 の所定期間である 100 ミリ秒遡及した時刻を示すインデックス  $i_b$  を次式により演算する。

$$\text{【0082】 } i_b \leftarrow (i-3) - 0.1 / \Delta t$$

30 なお、右辺第 2 項は 100 ミリ秒に存在するサンプリング回数である。ステップ 5e で音声信号が存在しない領域まで遡ることを防止するためにインデックス  $i_b$  が零以上に設定し、ステップ 5f で時刻を示すインデックスが  $i_b$  であるときの音声区間信号である  $g_2(i_b)$  を "1" に設定する。

【0083】ステップ 5g でインデックス  $i_b$  が処理時刻を示すインデックス  $(i-3)$  となったか、即ち第 2 の所定期間について遡及処理が完了したかを判定する。そして否定判定されたとき、即ち遡及処理が完了していないときはステップ 5h でインデックス  $i_b$  をデクレメントしてステップ 5f に戻る。逆に、ステップ 5g で肯定判定されたとき、即ち遡及処理が完了したときはこのルーチンを終了する。

【0084】図 24 は音声区間信号生成ルーチンのステップ 206 で実行される開処理ルーチンのフローチャートであって、ステップ 6a で前回演算されたゲート信号  $g_{1b}$  を "0" に設定する。ステップ 6b で閉継続時間  $t_{c0}$  を "0" に復帰し、ステップ 6c で処理時刻を示すインデックスが  $(i-3)$  であるときの音声区間信号である  $g_2(i-3)$  を "0" に設定してこのルーチを終了する。

50 【0085】図 25 は音声区間信号生成ルーチンのステ

ップ 207 で実行される第 2 のゲート信号出力ルーチンのフローチャートであって、ステップ 7a で処理時刻から第 2 の所定時間である 100 ミリ秒遡及した時刻を示すインデックス  $i_b$  を次式により演算する。 $i_b \leftarrow$

$$(i-3) - 0.1 / \Delta t$$

ステップ 7b で音声信号が存在しない領域まで遡ることを防止するためにインデックス  $i_b$  を零以上に設定し、ステップ 7c で  $g_2(i_b)$  を出力してこのルーチンを終了する。

【0086】図 26 は単語抽出部 28 で実行される単語抽出ルーチンのフローチャートであって、ステップ 260 で時刻を示すインデックス  $i_b$  であるときの単語信号  $W(i_b)$  を次式により算出する。

$$W(i_b) \leftarrow X(i_b) * g_2(i_b)$$

ただし  $X(i_b)$  は記憶部 24 に記憶されている音声信号である。

【0087】ステップ 261 で  $W(i_b)$  を出力してこのルーチンを終了する。

【0088】

【発明の効果】第 1 の発明に係る音声区間検出装置によれば、音声信号を時間領域で処理して抽出された音声ピッチに基づいてゲート信号が制御され、ゲート信号により音声区間が検出されるので、簡易な構成で音声区間を検出することが可能となる。第 2 の発明に係る音声区間検出装置によれば、音声区間に基づいて音声信号を複数の音声区間に区分することが可能となる。

【0089】第 3 の発明に係る音声区間検出装置によれば、音声信号を時間領域で処理して抽出された音声ピッチに基づいて音声区間が検出されるので、ほぼ実時間で音声区間を検出することが可能となる。第 4 の発明に係る音声区間検出装置によれば、音声信号の振幅の変動を抑制することが可能となる。

【0090】第 5 の発明に係る音声区間検出装置によれば、音声信号内に存在する雑音を確実に除去することが可能となる。第 6 の発明に係る音声区間検出装置によれば、音声信号の振幅を略一定に揃えることにより音声ピッチを確実に抽出することが可能となる。第 7 の発明に係る音声区間検出装置によれば、揃振幅利得が所定閾値であるときは揃振幅利得は単位利得に再設定することにより雑音の混入を防止することが可能となる。

【0091】第 8 の発明に係る音声区間検出装置によれば、ゲート信号が雑音の影響により誤って開となることを防止することが可能となる。第 9 の発明に係る音声区間検出装置によれば、ゲート信号が雑音の影響により誤って閉となることを防止することが可能となる。第 10 の発明に係る音声区間検出装置によれば、音声ピッチが抽出されなくなったときに第 1 ゲート信号を確実に閉とすることが可能となる。

【0092】第 11 の発明に係る音声区間検出装置によれば、ゲート信号の開遅れを補償するとともに、雑音を

気音と区別して確実に排除することが可能となる。第 12 の発明に係る音声区間検出装置によれば、振幅の小さい促音を確実に検出することが可能となる。第 13 の発明に係る音声区間検出装置によれば、音声区間が重複した場合にも誤検出を防止することが可能となる。

【図面の簡単な説明】

【図 1】従来のピッチ周期に基づく音声区間検出結果である。

【図 2】本発明に係る音声区間検出装置の機能構成図である。

【図 3】音声サンプリングルーチンのフローチャートである。

【図 4】前処理ルーチンのフローチャートである。

【図 5】ピッチ検出ルーチンのフローチャートである。

【図 6】減算処理ルーチンのフローチャートである。

【図 7】包絡線差算出ルーチンのフローチャートである。

【図 8】減算処理の効果の説明図である。

【図 9】AGC 処理ルーチンのフローチャートである。

【図 10】AGC の効果の説明図である。

【図 11】ピーク検出処理ルーチンのフローチャートである。

【図 12】極値検出・クランプ処理ルーチンのフローチャートである。

【図 13】ピッチ周期検出処理ルーチンのフローチャートである。

【図 14】ピッチ周期検出方法の説明図 (1/2) である。

【図 15】ピッチ周期検出方法の説明図 (2/2) である。

【図 16】第 1 のゲート信号生成ルーチンのフローチャートである。

【図 17】ゲート信号の生成方法説明図である。

【図 18】音声信号処理例である。

【図 19】第 2 のゲート信号生成ルーチンのフローチャートである。

【図 20】音声区間信号生成ルーチンのフローチャートである。

【図 21】閉維持処理ルーチンのフローチャートである。

【図 22】開処理ルーチンのフローチャートである。

【図 23】開維持処理ルーチンのフローチャートである。

【図 24】閉処理ルーチンのフローチャートである。

【図 25】音声区間信号出力ルーチンのフローチャートである。

【図 26】単語抽出ルーチンのフローチャートである。

【符号の説明】

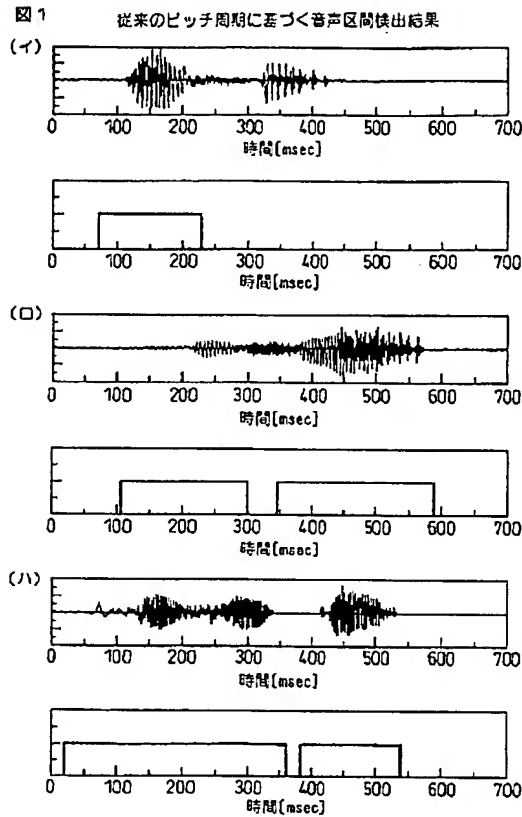
21…マイクロホン

22…ライン増幅器

23…アナログ／デジタル変換部  
24…記憶部  
25…ピッチ検出部

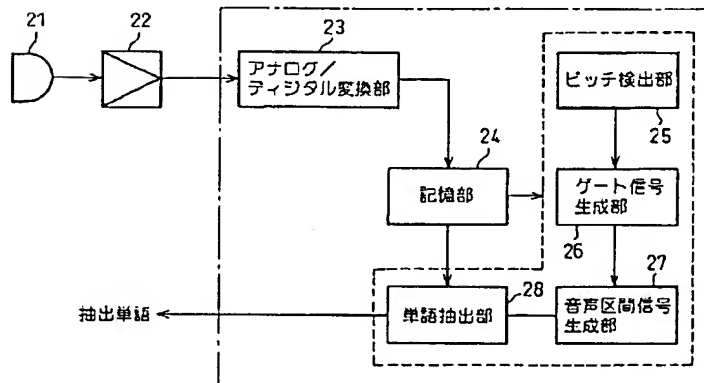
26…ゲート信号生成部  
27…音声区間信号生成部  
8…単語抽出部

【図1】



【図2】

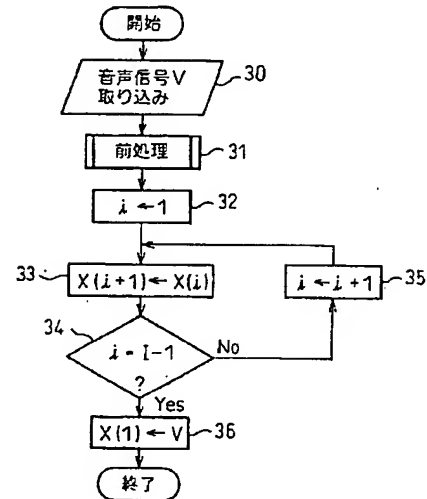
本発明に係る音声区間検出装置の機能構成図



【図3】

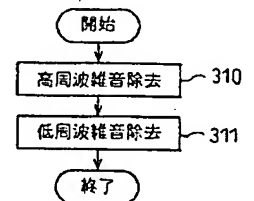
図3

音声サンプリングルーチンのフローチャート



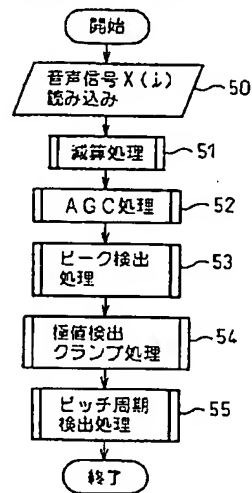
【図4】

図4 前処理ルーチンの詳細フローチャート



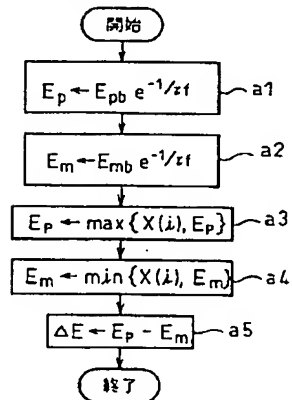
【図5】

図5 ピッチ検出ルーチンのフローチャート



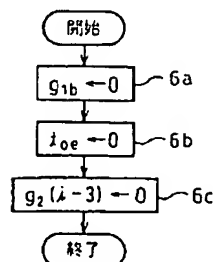
【図7】

図7 包値差算出ルーチンのフローチャート



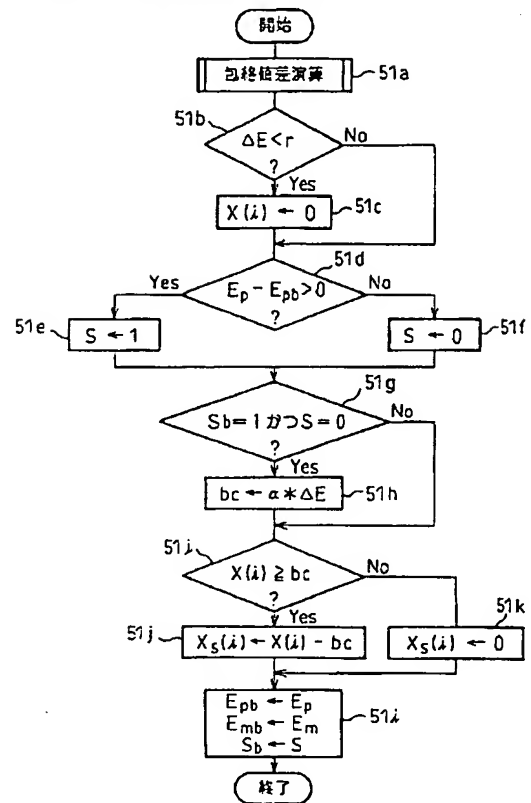
【図24】

図24 閉処理ルーチンのフローチャート



【図6】

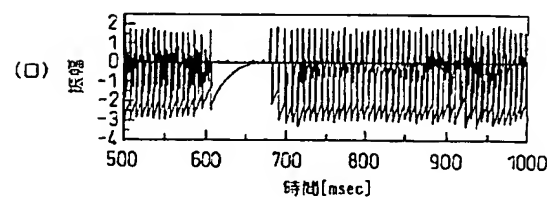
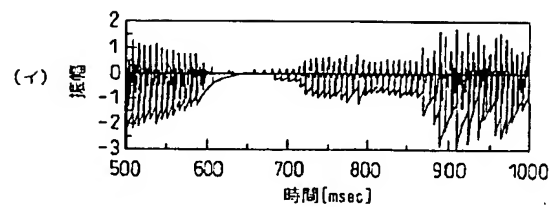
図6 減算処理ルーチンのフローチャート



【図10】

図10

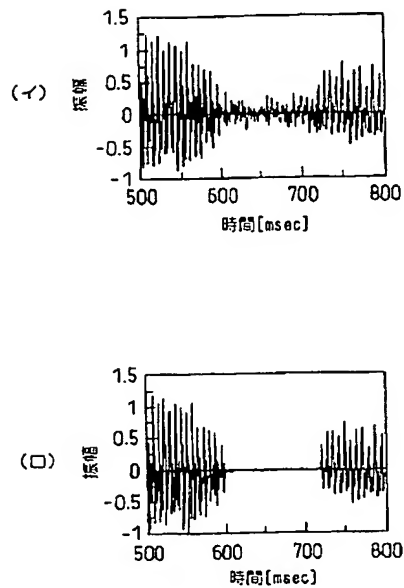
AGCの効果の説明図



【図8】

図8

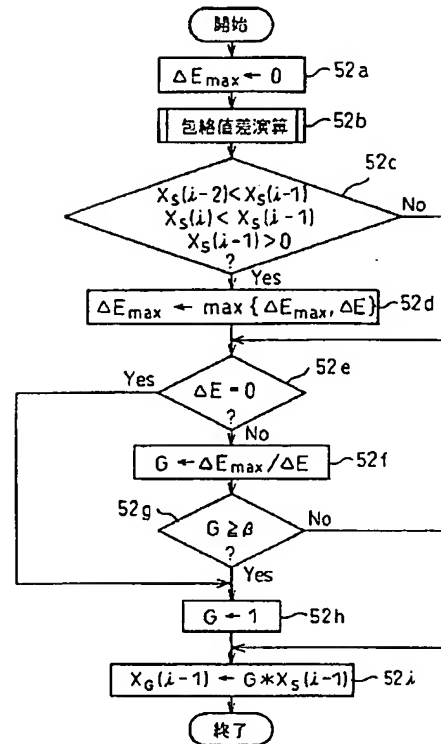
減算処理の効果の説明図



【図9】

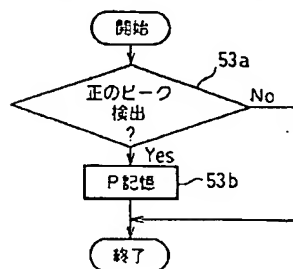
図9

AGC処理ルーチンのフローチャート



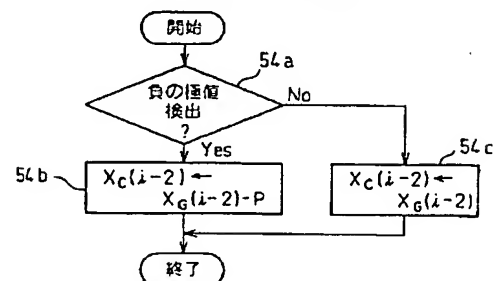
【図11】

図11 ピーク検出処理ルーチンの詳細フローチャート



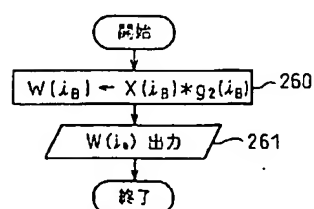
【図12】

図12 極値検出・クランプ処理ルーチンの詳細フローチャート



【図26】

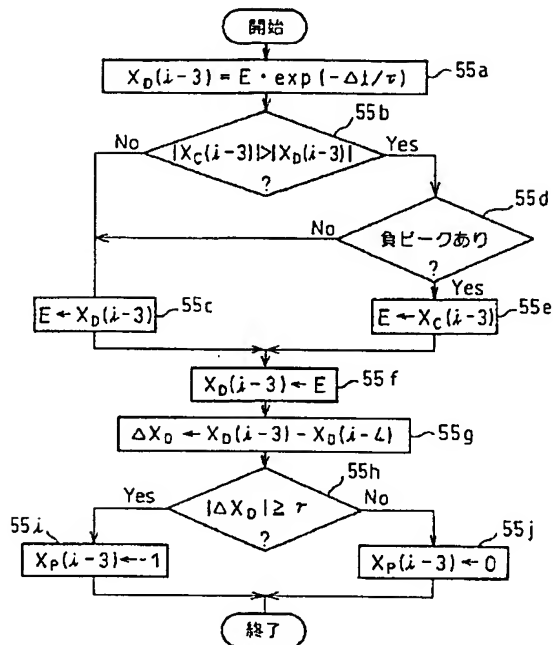
図26 単語抽出ルーチンのフローチャート



【図13】

図13

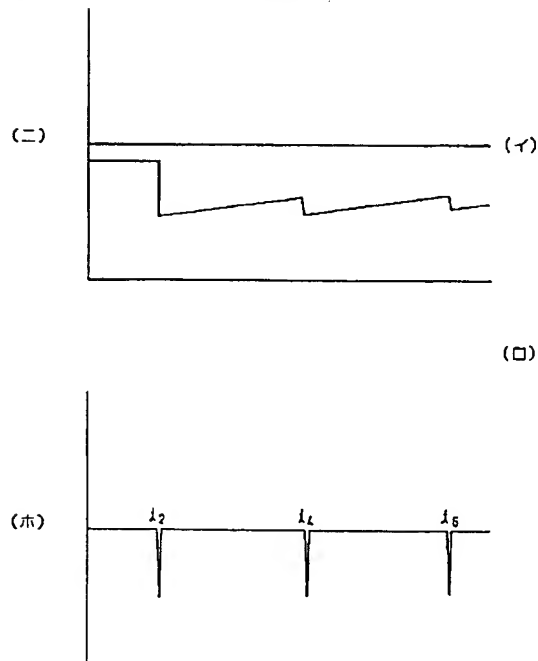
ピッチ周期検出処理ルーチンの詳細フローチャート



【図15】

図15

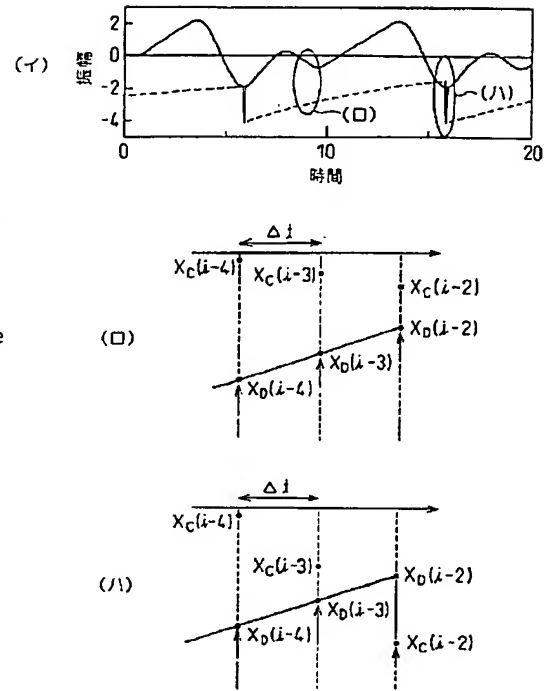
ピッチ周期検出方法の説明図(2/2)



【図14】

図14

ピッチ周期検出方法の説明図(1/2)



【図17】

ゲート信号の生成方法説明図

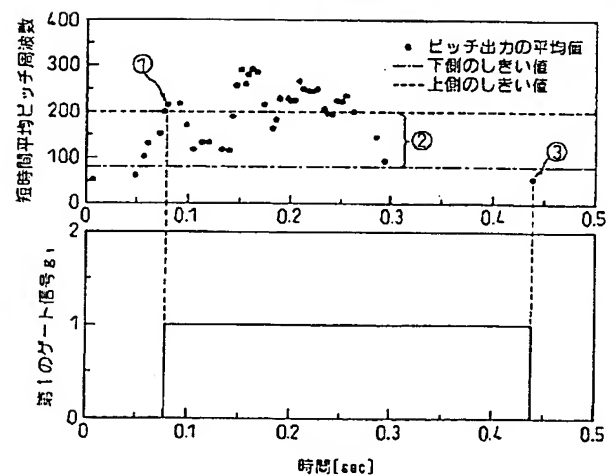
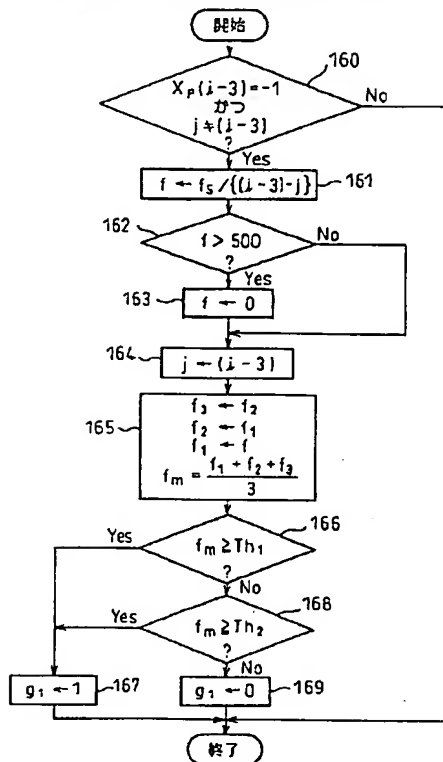


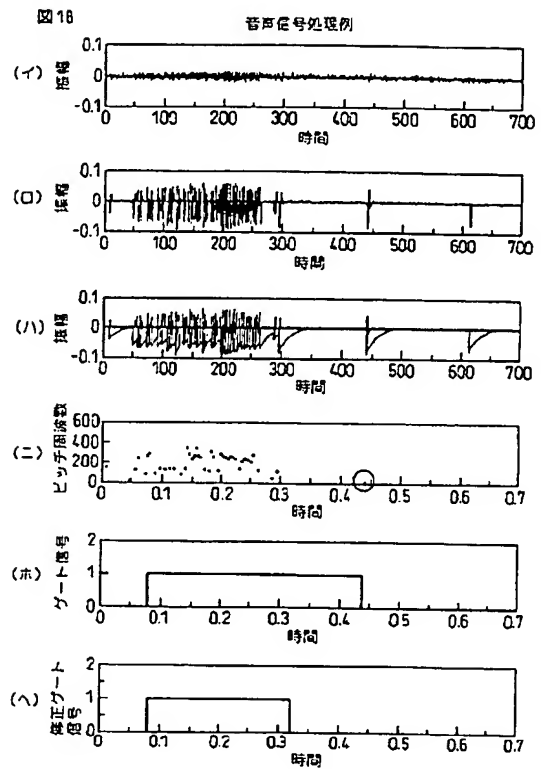
図17

【図16】

図16 第1のゲート信号生成ルーチンのフローチャート

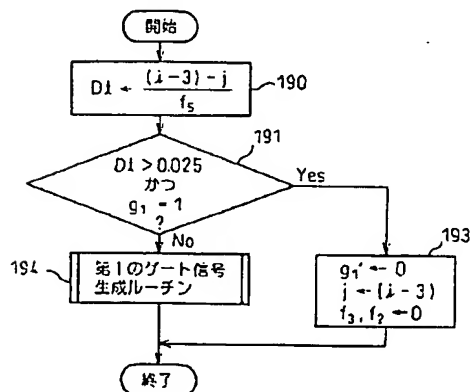


【図18】



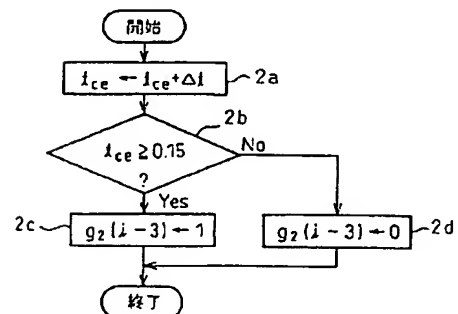
【図19】

図19 第2のゲート信号生成ルーチンのフローチャート



【図21】

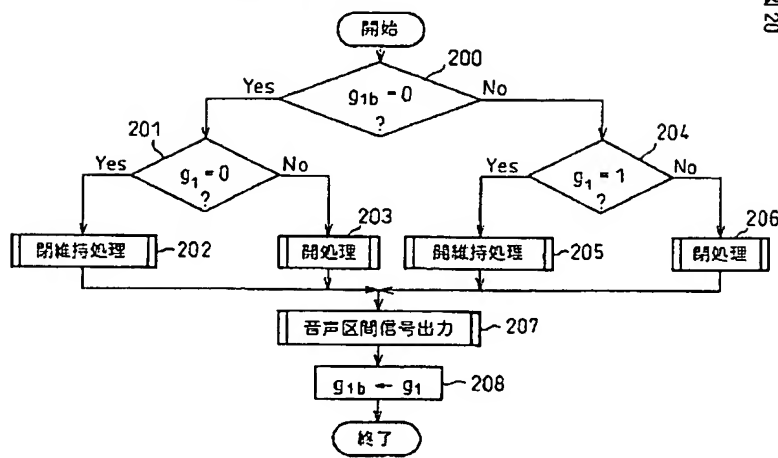
図21 閉鎖維持処理ルーチンのフローチャート





【図20】

音声区間信号生成ルーチンのフローチャート

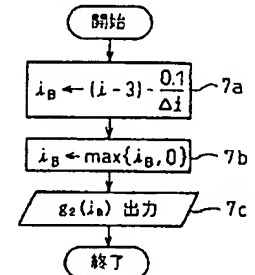


【図25】

図25

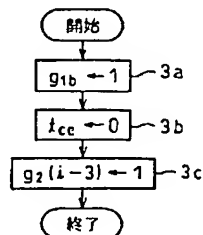
音声区間信号出力ルーチンのフローチャート

図20



【図22】

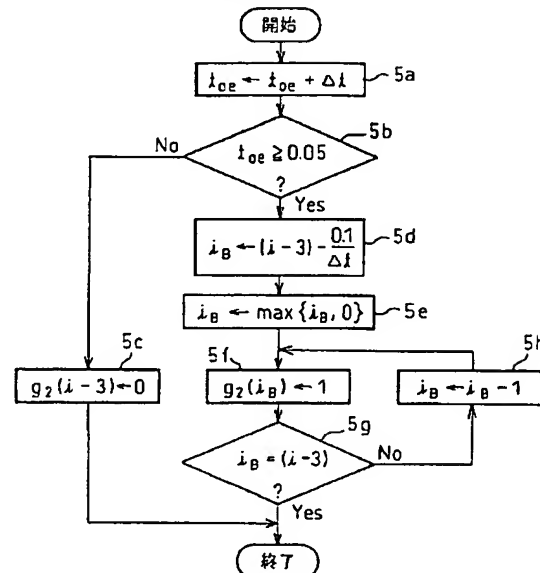
図22 閉処理ルーチンのフローチャート



【図23】

図23

閉維持処理ルーチンのフローチャート



フロントページの続き

(72)発明者 北尾 英樹  
兵庫県神戸市兵庫区御所通1丁目2番28号  
富士通テン株式会社内

(72)発明者 岩本 真一  
兵庫県神戸市兵庫区御所通1丁目2番28号  
富士通テン株式会社内

(72)発明者 岩田 収  
兵庫県神戸市兵庫区御所通1丁目2番28号  
富士通テン株式会社内

(72)発明者 中村 正孝  
広島県広島市佐伯区三宅二丁目1-1 学  
校法人鶴学園内

(17)

特開 2 0 0 2 - 9 1 4 7 0

(72) 発明者 大元 芳尚  
東京都杉並区久我山 1 - 5 - 25

F ターム ( 参考 ) 5D015 AA05 DD03 EE05